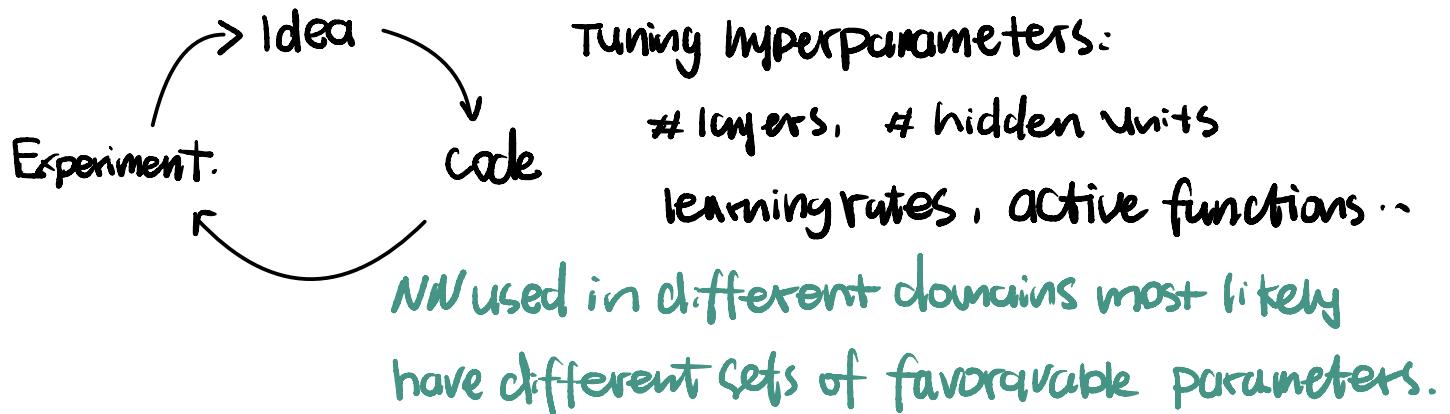


Train / dev / test sets.

* Applied ML is a highly iterative process.



Train / dev / test sets.

Data

	dev	test
--	-----	------

training data.

↑

Prev-era: 70 / 30, 60 / 20 / 20

- Hold-out cross validation

Big-data : 1M, 1K, 1K

- Development set. "dev"

→ 99.5%, 0.4%, 0.1% ~

We typically don't need that much data for dev and test, so in big-data era, much higher proportion of data is used in training

* Mismatched train/test distribution

e.g. Training set :

Cat pictures from
web pages.

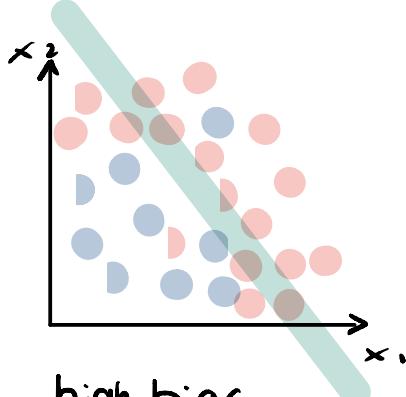
(high resolution, clear)

Dev / test set :

Cat pictures from users
using your app
(unclear, casual)

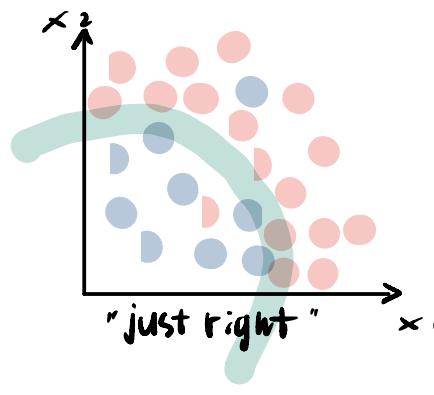
Solution : Make sure at least : dev and test come from same distribution. not having test might be strong (+train / dev).

Bias / Variance

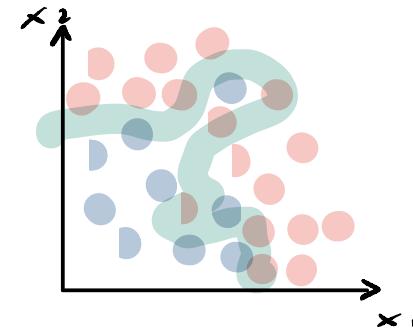


high bias

underfit: not doing good
at fitting training set.



"just right"



high Variance

overfit: fit too much
over training data. Not
generalized. Won't perform
good on test data.

Bias and Variance: high bias.

Training Error: 1%

Dev set Error: 11%

high variance

15%

16%

15%

30%

0.5%

1%

low bias

low variance

high bias

high variance.

With the assumption:

Optimal (Bayes) error $\approx 0\%$.

This can be a low bias and low variance
network if the bayes error is 15%
e.g. classifying blurry image.

high bias =

Poor performance on
training data

high Variance:

Poor Performance on testing data.

BASIC Recipe for machine Learning.

